

## NetApp、NEC、富士フィルムが 三位一体で、 IoTやデータ利活用推進を支援



NetApp INSIGHT Japan 2021に、富士フィルム株式会社と日本電気株式会社が共同で「次世代技術開発におけるデータ利活用、結果に差が出るAI・IoT基盤とは?」と題する講演をオンラインで開催しました。富士フィルム株式会社の記録メディア事業部 シニアエンジニアの森純也氏は「みなさまが感じている課題解決のヒントになれば幸いです」と切り出し、AIやIoT基盤におけるデータ利活用の課題と解決のヒントについて講演しました。



富士フィルム株式会社  
**森 純也**  
記録メディア事業部  
シニアエンジニア



日本電気株式会社  
**大須 隆寿**  
システムプラットフォーム  
ビジネスユニット  
AIプラットフォーム事業部  
データ活用基盤担当

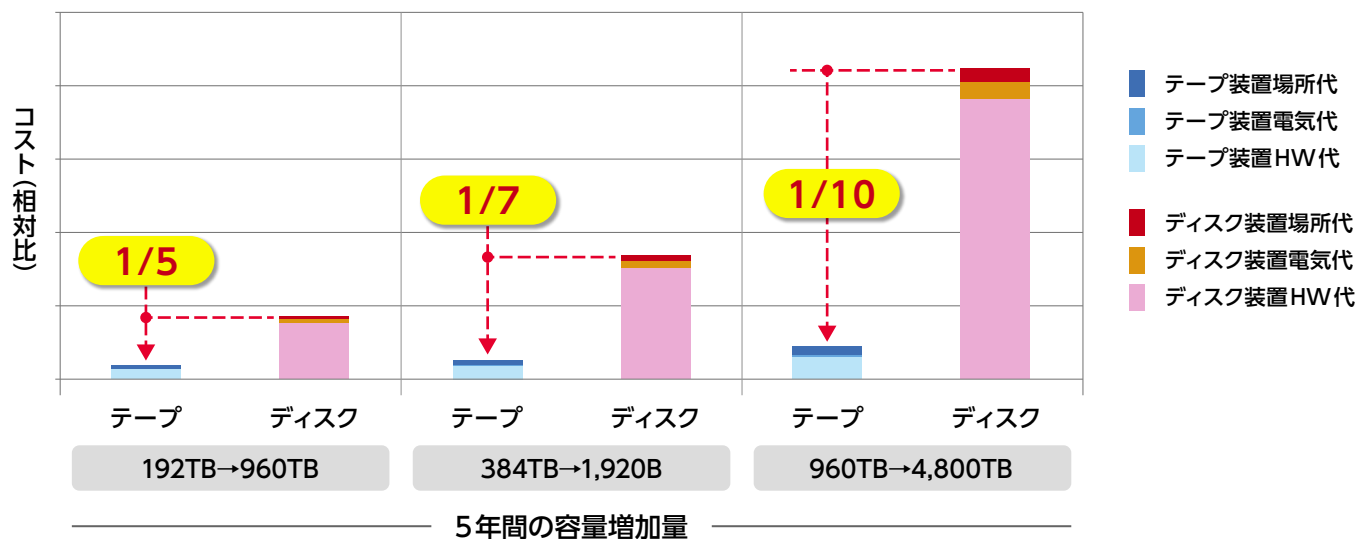
### データ利活用における課題と富士フィルムの解決提案

ヒトゲノムや医薬から自動運転にIT建機、そして宇宙や深海探査に至るまで、研究や開発に分析といった分野で、非構造化データをはじめとした膨大なデータの利活用が進んでいます。こうした取り組みが進む中で、森氏は「データ利活用のための大容量データ保管が必要になり、保管のためのストレージコストが障壁になっています」と指摘します。分析のようにデータを「使う」ためのストレージは、高パフォーマンスのデバイスが利用されていますが「溜める」ためのメディアとして、テープストレージの利用が注目されています。森氏はテープストレージの良さについて「非常に安価で、エアギャップの発生をメリットにして、BCP対策やランサムウェア対策などの効果が得られます。テープストレージのエラー率はHDDの1/10000と低く、メディアは50年という長期保管にも対応しています」と解説します。大容量で低コストのテープストレージは、現在は12TBが主流となっていますが、研究所による試験では1つのカートリッジで580TBまで保管できる将来性もあります。

また、HDDと比較すると保存するデータ容量が多くなるほど、テープストレージの優れたコストパフォーマンスが確かめられます。森氏は「HDDでは、保存容量を増やすためには、ディスク装置をどんどん積まないといけません。スケールアウトによるHDDの増設は、電気代もハウジングコストもかかります。それに対してテープストレージであれば、メディアを買い増していくだけで容量を増やせるので、低コストで大容量に対応できます」と説明します。

さらに、森氏は「シーケンシャルな書き込みでは、テープは非常に高速です。LTOで360MB/secの書き込みが可能で、その実績が高く評価され、欧州原子核研究機構 (CERN) では、大型ハドロン衝突型加速器によって得られた実験の未加工データを保存する目的で、テープストレージを採用しました。また、日本の高エネルギー加速器研究機構 (KEK) でも、容量単位の安さと、万が一のメディア不良でも、すべてのファイルが読めなくなることがない、というサルベージできる安心感からテープストレージを選ばれました」と話します。

## ■ テープストレージは大容量で低コスト



※テープ装置：80巻テープライブラリ、LTO 8ドライブ搭載 (非圧縮 12TB)

※ディスク装置：RAID 6構成、高密度実装タイプ、エコモード、Near Line 12TB HDD

## 大容量データの保管コストを最適化するソリューション

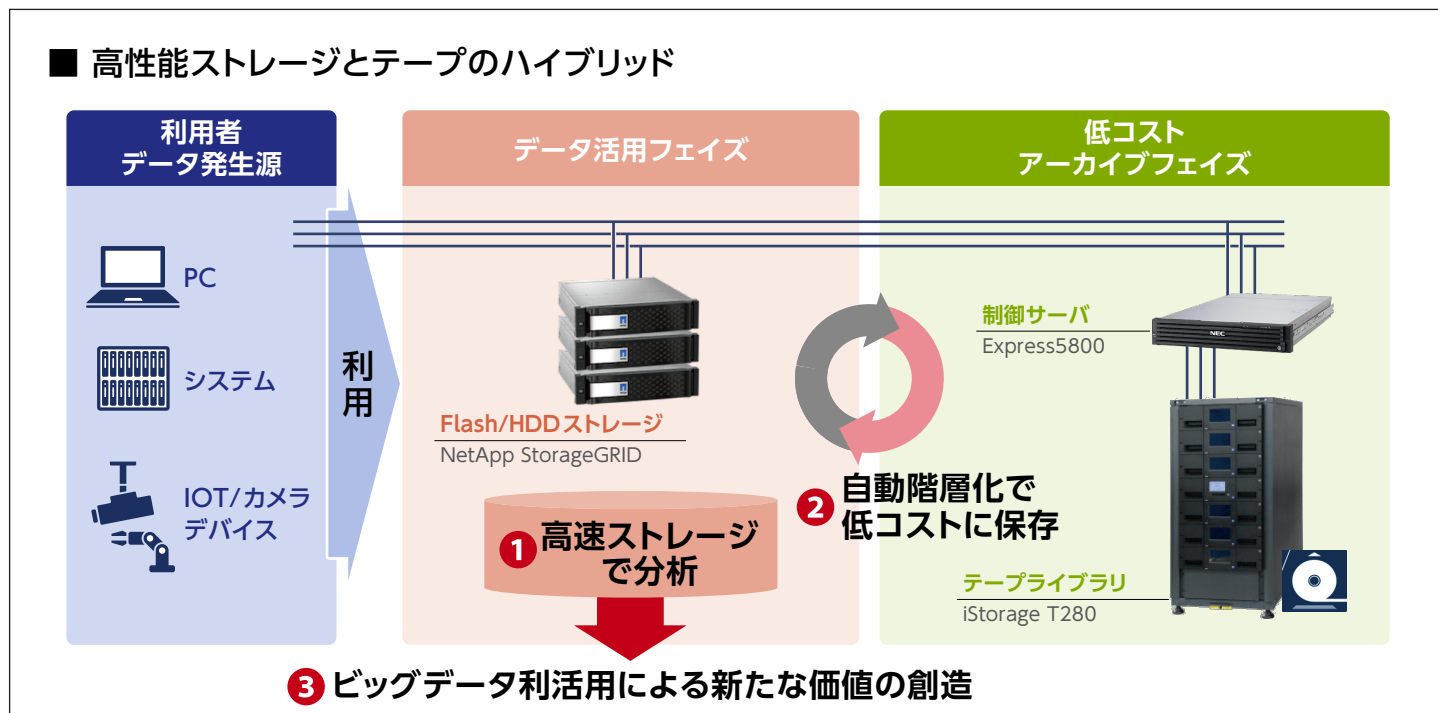
大容量データの保管コストを最適化するソリューションとして、富士フィルムはオブジェクトストレージの利用を提案します。オブジェクトストレージは、ファイルをオブジェクトとして扱い、データ本体をIDとメタデータで管理するストレージ技術です。Windowsなどで一般的に利用されているファイルストレージに比べて、大容量のデータ保管に適したストレージソリューションとなります。オブジェクトストレージでは、階層構造を持たずにフラットにデータを扱います。HTTPベースで使いたいデータを使いたいように使える利点があります。

森氏は「オブジェクトストレージに対応したFUJIFILM オブジェクト アーカイブは、S3互換APIを実装し、オブジェクト形式でテープにデータを読み書きするソフトウェアです。FUJIFILM オブジェクト アーカイブをご活用いただくと、オンプレミスに安価で堅牢な大容量オブジェクトストレージを構築できます」と説明します。

大容量のオブジェクトデータを長期間保存するときに、FUJIFILM オブジェクト アーカイブを使うと、オンプレミスに『Amazon AWS、GlacierとLife Cycle機能』を再現することが可能なのです。森氏は「機密データやパブリッククラウドに置けない情報や、AWSで長期に保管したいデータなどにオブジェクトアーカイブを使うと、簡単に安価に対応できます」と提案し、高性能ストレージとテープをハイブリッドに組み合わせたストレージソリューションを提案します。

そして、高性能ストレージとテープストレージをハイブリッドに組み合わせた、ビッグデータ利活用の新たな価値想像の取り組みを紹介します。

森氏は「左側は、利用者によるデータの発生源を示します。これらのデータを先ずはNetAppのストレージグリッドに保存して、高速にデータ分析などを行います。その分析基盤には、この後でお話するNECのソリューションを参考にしてください。そして、データの活用が終わった後に、自動的にテープへGlacierのストレージクラスとして保存することで、ビッグデータ利活用の新たな価値創造につながります」と説明します。



## 分析を高速化するNECのソリューション

続いて、日本電気株式会社 システムプラットフォーム ビジネスユニット AIプラットフォーム事業部 データ活用基盤担当 大須隆寿氏が、高性能ストレージとテープのハイブリッド環境を活かすNECのソリューションを紹介しました。大須氏は「データ活用は、5つのステージで構成されています。BAやAIを活用しようと検討していても、この5つのステージを着実に進んでいかないと、データ活用は促進できません」と説明します。

データ活用を見据えたデータ連携について、大須氏は「単にデータを連携するのではなく、現状を振り返り、今後のあるべき姿を見据えてデータを連携する必要があります。そのポイントは、目的に合わせたデータ統合に、クレンジングや加工によるデータの品質の担保、そしてGUIによる変更や管理の簡易化です」と解説します。



## ■ 格納 ～用途にあったデータ格納先～

	データレイク	データウェアハウス	データマート
	多種多様なデータを元のフォーマットのまま保管するデータ格納領域	基幹系システムなどからデータを収集し、時系列にデータを保管	利用部門や用途・目的に応じて、データを抽出・加工・集計し、利用しやすい形で格納
格納データの種類	非構造化データ・半構造化データ・構造化データ	構造化データ	構造化データ
データ保持形式	ソースから取得したデータを加工せずのそのまま保持	時系列のトランザクションデータ	目的にあわせて加工されたサマリデータ
データ保持期間	長い ←		→ 短い
データ量	数十TB～数PB以上	数百GB～数PB	数十GB程度

データの連携・格納先は3つあり、データレイクやデータウェアハウス、データマートと呼ばれている3つのデータ格納先において大須氏は魚を例にして、多種多様なデータが保管されているデータレイクという格納領域を湖に例えます。その湖から釣り上げられた魚が保管される倉庫がデータウェアハウスになります。そして、用途や目的に応じて加工されるデータマートは、切り身などが並ぶ販売店にあたと例えます。大須氏は「これらデータ格納の場におけるデータレイクは、なるべく安価なストレージの採用が、重要なポイントになります」と補足します。さらに大須氏は「NECはデータ活用基盤に対して、最適なプラットフォームを一気通貫してご提案させていただきます。お客様のビジネスや用途の拡大に伴うデータ活用基盤の拡張性を確保し、連携・格納・分析までの一気通貫したご提案により、製品の組み合わせによる相乗効果や実現可能な課題解決策まで検討し、構築後も継続したご支援が可能です」と提案します。

講演の最後に、富士フィルムの森氏は「NetApp様、NEC様、富士フィルムが三位一体となって、皆様のIoTやデータ利活用推進をご支援します」と結びました。

